International Journal of Scholarly Research (IJSR)
Vol-1, Issue-2, 2017
http://www.scholars.org.in

"Implementation of Efficient Approach towards Classification of Semantically Secure Encrypted Data"

Gauravweni Hedau[1], Prof.Vikrant Chole[2]

[1] Departement of computer science and engg,
G.H. Raisoni Academy of engg. & technology, Nagpur
ghedau1373@gmail.com

[2] Departement of computer science and engg,
G.H. Raisoni Academy of engg. & technology, Nagpur
Vikrant.chole@raisoni.net

Abstract*:* As per our review paper we had overview the various related work of classification of semantically secure encrypted data. More than that we also overview the proposed working of our new system which we are going to devolve. Now, in this proposed paper we are going to discuss implementation detail of this proposed system. As we have clear that previously we are using secure k-NN for classification over encrypted data but this work is further extended to provide a new solution. In our paper we are proposing privacy preserving k-nn (PPk-NN) to achieve classification over encrypted data. When we increase or enhance the scope of data from local server to cloud that is in distributed environment so before classification we have to encrypt data that means convert plain text into cipher text. As we know data that is going to distribute on cloud is already at some server and it is already in encrypted form done by its provider or the data owner. We are recommending a homomorphic encryption method to solve this issue of encrypting already encrypted data before classification.

Keyword: Outsourced database, PPk-NN, homomorphic encryption, cloud, security.

1. INTRODUCTION

Cloud computing is a budding computing paradigm in which resources of the computing infrastructure are provided as services over the Internet. In this budding cloud computing paradigm there will be a more security issues rises in the world today. So we are proposing the system that allows us more security in terms of data privacy. As promising as it is, cloud computing is also facing many challenges that, if not well resolute, may slow down its fast growth.

There are many data security for privacy purpose application are available in the real world, but in cloud computing system these dispute would great concerns from users when they store sensitive information on cloud servers. We have to put more interest on this fact of data security because this cloud server is used or accomplished by commercial providers which are very probable to be outside of the trusted domain of the users.

Data privacy or confidentiality against cloud servers is hence repeatedly preferred when users outsource data for storage in the cloud. When data are exceptionally susceptible and we do not want it to share with anyone who is not credible party, so for that the data need to be encrypted before outsourcing to the cloud. However, when data are encrypted, irrespective of the fundamental encryption scheme, performing any data mining tasks becomes very challenging without ever decrypting the data the privacy/security requirements of the DMED(data mining over encrypted data) problem on a cloud are threefold:

(1) Confidentiality of the encrypted data,
(2) Confidentiality of a user's query record, and
(3) Hiding data access patterns.

Methods to successfully solve the DMED problem assuming that, the encrypted data are outsourced to a cloud.

Specifically, we focus on the classification problem since it is one of the most common data mining tasks. There are many classification approach are available, because each classification approach has their own plus points, to be tangible, in this paper we intensify on assassinate the PPk-NN classification method over encrypted data in the cloud computing surroundings. The information inquiry record has to be encrypted before transmitting it to the cloud because it will contain sensitive information.

We are going to discuss in next section all the details in the proposed working flow. In section 2 we are going to discuss literature work. In next Section 3 we are going to review related work. Then in section 4 we are going to discuss detail working with flowchart and architecture that means proposed working. In section 5 we will describe experimental result and execution process. And at the end in section 6 we give the conclusion.

2. LITERATURE WORK

In this section we are discuss previously work done by some extraordinary developer who have worked on the Classification of Semantically Secure Encrypted Data.

S. De Capitani di Vimercati, S. Foresti and P. Samarati [1]. They have point up risks, solutions, and open problems related

to insure privacy of users invoking services or resources in the cloud, sensitive private information stored at un-trusted domain, and accesses to such information.

B. K. Samanthula, Y. Elmehdwi and W. Jiang [2] have deliberated on solving the classification problem over encrypted data. They have projected a secure k-NN classifier over encrypted data in the cloud computing environment.

C. Gentry and S. Halevi's [3] main expansion is a key-generation process for the core somewhat homomorphic encryption, that does not require full polynomial inversion.

Yehuda Lindell and Benny Pinkasy [4] have proposed basic expansions and ideas of protected secured multiparty computation and discuss their significance to the field of privacy-preserving data mining (PPDM).

H. Hu, J. Xu, C. Ren and B. Choi [5] have proposed a comprehensive and competent solution that comprises a secure traversal framework and an encryption scheme based on privacy homomorphism.

Apurva Gomase, Prof.Vikrant Chole [6] has projected re-encryption in which the data is encrypting twice. So this technique is efficient and extensible to securely handle users private and sensitive that do not want to exposes data in the data sharing system user ensures about the data storage in external data storing center.

Somesh P. Badhel, Prof. Vikrant Chole [7] have given the review on all the techniques and tried to cover different issues of data backup and recovery of data after damage for Cloud Computing such as maintaining the cost of implementation and implementation complexities as low as possible.

Somesh P. Badhel, Prof. Vikrant Chole [8] have obtainable feature design of projected Backup recovery technique for cloud computing.

3. PROPOSED METHOD

We have studied the different system design from the literature work as describe in the previous section. Our proposed system design goal is to perform the classification on already encrypted data when the customer queries the cloud server. For user privacy, there are many privacy preserving mechanisms are available. So we use PPk-NN classification mechanism for classification on already encrypted data for user confidentiality and hide data access patterns. PPk-NN is a more multifarious problem and it cannot be solved directly or easily using the existing secure k-nearest neighbor mechanism over encrypted data. The goal of PPk-NN procedure is to classify (extract) users query records or required data qurey using privacy preserving technique. The PPk-NN protocol mainly consists of the following two stages:

- Stage 1 - Secure Retrieval of k-Nearest Neighbors (SRkNN)

In this stage, an authorized authenticated user primarily sends his query q (in encrypted form) to C1. After this, C1 and C2 occupy in a set of sub-protocols to steadily repossess (in encrypted form) the class labels corresponding to the k-nearest neighbors of the input query q. At the end of this step, encrypted class labels of k-nearest neighbors are known only to C1 that is cloud1.

- Stage 2 - Secure Computation of Majority Class (SCMCk)

Following from Stage 1, C1 and C2 jointly calculate the class label with a preponderance voting among the k-nearest neighbors of q. At the end of this step, only approved authorized user knows the class label equivalent to his input query record q.

The outcome data of Stage 1 which are passed as input to Stage 2 are in encrypted format. Therefore, the chronological composition of the two stages leads to our PPk-NN mechanism and we claim it to be secure underneath the semi-honest model according to the Composition Theorem.

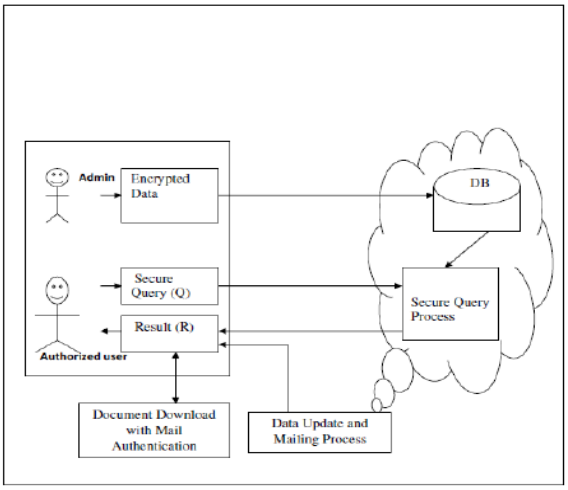

Fig. 1. Architecture Diagram for PPk-NN

Homomorphic encryption allows difficult mathematical operations to be performed on encrypted data without using the original data and provides the data security and confidentiality in cloud. The proposed PPk-NN protocol is used for classification mechanism usually applied for data mining task for extracting the required data. It determines which the nearest results are by identifying the class of bare minimum distance using K nearest neighbors. Refer figure 3.1 to privacy preservation for data in cloud.

In our paper we proposed PPk-NN protocol which is privacy preserving protocol useful over data which is in previously encrypted form by its provider data owner and then we are going to apply homomorphic encryption technique on this cipher text data , because homomorphic encryption have functionality to encrypt data which is in already in encrypted form means in cipher text. That mean we apply encryption on cipher text rather than plain text which solve our main problem and we will accomplish our main goal.

For this scenario we have probably three steps:

A. Data upload which is in encrypted form by data owner.

B. Apply homomorphic encryption on already encrypted data.
C. PPk-NN for query processing query on cloud environment.

In our proposed work we have to encrypt the data which is available at the data owner side. When the query held by the customer the classification is performed over the encrypted data which is decrypted with the corresponding decryption key which is available at the another cloud server.

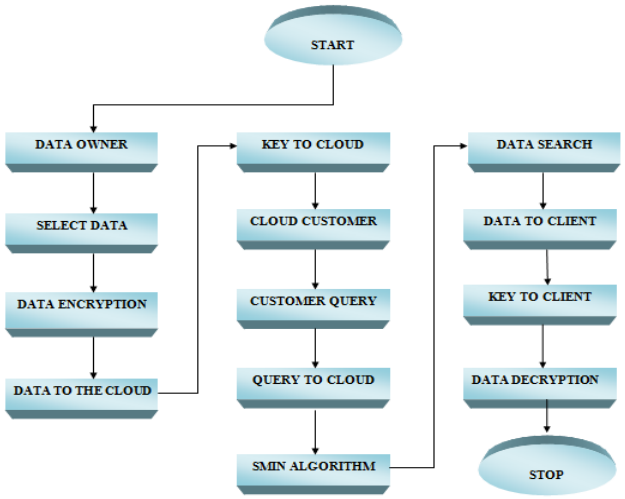

Fig. 2. Flowchart

Workflow of system:

1. Data owner is an important person. Data owner have to select data his data is to be encrypted and stored in cloud1. The key used for to encrypt the data is stored on cloud2. There are two cloud in the system cloud1 is use to store encrypted data and cloud2 is use to store encryption key.
2. Then the customer enters into mechanism means in a system. Then customer sends a query to the server that is a cloud 1.
3. Cloud 1 searches the relevant data for the query send through the customer by the SMIN algorithm. This algorithm is used to search the relevant data to the query send to by the customer.
4. Then the relevant data are finding by the cloud1 and transferred to the client. This data is in the form of encryption and then client get the key from the cloud2. The encrypted data from the cloud 1 is decrypted then and only then the client is able to view the original data.

Algorithms:

1. SMIN (Secure Minimum):

Algorithm 1 SMIN(u′, v′) → [min(u, v)],Epk(smin(u,v))

Require: P1 has u′ = ([u],Epk(su)) and v′ = ([v],Epk(sv)), where 0 ≤ u, v < 2l; P2 has sk

1: P1:

(a). Randomly choose the functionality F

(b). **for** i = 1 to l **do**:

• Epk(ui ∗ vi) ← SM(Epk(ui),Epk(vi))

• Ti ← Epk(ui ⊕vi)

• Hi ← Hri

i−1 ∗ Ti; ri ∈ R ZN and H0 = Epk(0)

• _i ← Epk(−1) ∗ Hi

• **if** F : u > v **then**:

– Wi ← Epk(ui) ∗ Epk(ui ∗ vi)N−1

– □i ← Epk(vi − ui) ∗ Epk(ˆri); ˆri ∈ R ZN

else

– Wi ← Epk(vi) ∗ Epk(ui ∗ vi)N−1

– □i ← Epk(ui − vi) ∗ Epk(ˆri); ˆri ∈ R ZN

• Li ← Wi ∗ _r′

i

i ; r′

i ∈ R ZN

(c). **if** F : u > v **then**: _ ← Epk(sv − su) ∗ Epk(¯r)

else _ ← Epk(su−sv)∗ Epk(¯r), where ¯r ∈ R ZN

(d). □′ ← _1(□) and L′ ← _2(L)

(e). Send _, □′ and L′ to P2

2: P2:

(a). Receive _, □′ and L′ from P1

(b). Decryption: Mi ← Dsk(L′

i), for 1 ≤ i ≤ l

(c). **if** ∃ j such that Mj = 1 **then** _ ← 1

else _ ← 0

(d). **if** _ = 0 **then**:

• M′

i ← Epk(0), for 1 ≤ i ≤ l

• _′ ← Epk(0)

else

• M′

i ← □′

i ∗ rN, where r ∈ R ZN and is

different for 1 ≤ i ≤ l

• _′ ← _ ∗ rN

_ , where r_ ∈ R ZN

(e). Send M′,Epk(_) and _′ to P1

3: P1:

(a). Receive M′,Epk(_) and _′ from P2

(b). fM ← _−1

1 (M′) and _ ← _′ ∗ Epk(_)N−¯r

(c). _i ← fMi ∗ Epk(_)N−ˆri , for 1 ≤ i ≤ l

(d). **if** F : u > v **then**:

• Epk(smin(u,v)) ← Epk(su) ∗ _

• Epk(min(u, v)i) ← Epk(ui)∗ _i, for 1 ≤ i ≤ l

else

• Epk(smin(u,v)) ← Epk(sv) ∗ _

• Epk(min(u, v)i) ← Epk(vi)∗ _i, for 1 ≤ i ≤ l

2. Homomorphic Encryption algorithm:

- Encryptions that allow computations on the ciphertexts
  - $E_k[m_1] \bullet E_k[m_2] = E_k[m_1 \circ m_2]$
- Applications
  - E-voting: everyone encrypts votes as 1 or 0, aggregate all ciphertexts before decrypting; no individual vote is revealed.
    - Requires additive homomorphic encryption: ° is +
  - Secure cloud computing.
    - Requires full homomorphic encryption, i.e., homomorphic properties for both + and ×

Homomorphic Properties of Some Encryption Schemes

- Multiplicative homomorphic encryption
  - Unpadded RSA: $m_1^e \times m_2^e = (m_1 \times m_2)^e$
  - El Gamal: Given public key $(g, h=g^a)$, ciphertexts $(g^{r1}, h^{r1}m_1)$ and $(g^{r2}, h^{r2}m_2)$, multiple both components $(g^{r1+r2}, h^{r1+r2}m_1m_2)$
- Additive homomorphic encryption schemes
  - Paillier cryptosystem (will explore in HW problem)
- Fully homomorphic encryption also exist
  - Significantly slower than other PK encryption

4. CONCLUSIONS

In Cloud computing environment providing the security and confidentiality to the third party customer is main issue is solved in this proposed system. To secure the customer query processing and data transferring to the customer as per queried through the user all this issues are solved in our proposed system through the PPK-NN protocol. And performing encryption over encrypted data is performed by the homomorphic encryption mechanism.

4. REFERENCES

1. S. De Capitani di Vimercati, S. Foresti and P. Samarati, "Managing and accessing data in the cloud: Privacy risks and approaches", Proc. 7th Int. Conf. Risk Security Internet Syst., pp. 1-9, 2012.
2. B. K. Samanthula, Y. Elmehdwi and W. Jiang, "k-nearest neighbor classification over semantically secure encrypted relational data", 2014.
3. Gentry, "Fully homomorphic encryption using ideal lattices,"Proc. 41st Annu. ACM Sympos. Theory Comput., pp. 169-178, 2009.
4. Y. Lindell and B. Pinkasy, "Privacy preserving data mining", Proc. 20th Annu. Int. Cryptol. Conf. Adv. Cryptol., pp. 36-54, 2000.
5. Haibo Hu, Jianliang Xu, Chushi Ren, Byron Choi," Processing private queries over untrusted data cloud through privacy homomorphism", IEEE Xplore, May 2011.
6. Apurva Gomase, Prof.Vikrant Chole, "Secure system implementation using attribute based encryption", ijates,Vol.No.03,Special issue No.01,Nov 2015.
7. Somesh P. Badhel, *vikrant Chole*, "A review on data back-up techniques for cloud computing", International Journal of Computer Science and Mobile Computing, Vol.3 Issue.12, December- 2014, pg. 538-542.
8. Somesh P. Badhel, Prof. Vikrant Chole, "An efficient and secure remote data back-up technique for cloud computing", International Journal of Computer Science and Mobile Computing, Vol.4 Issue.6, June-2015, pg. 361-36.
9. P. Mell and T. Grance, "The NIST definition of cloud computing (draft)", NIST Special Publication, vol. 800, pp. 145, 2011.
10. P. Williams, R. Sion and B. Carbunar, "Building castles out of mud: Practical access pattern privacy and correctness on untrusted storage,"Proc. 15th ACM Conf. Comput. Commun. Security, pp. 139-148, 2008.
11. P. Paillier, "Public key cryptosystems based on composite degree residuosity classes,"Proc. 17th Int. Conf. Theory Appl. Cryptographic Techn., pp. 223-238, 1999.
12. Gentry and S. Halevi, "Implementing gentry¿¿¿s fully-homomorphic encryption scheme,"Proc. 30th Annu. Int. Conf. Theory Appl. Cryptographic Techn.: Adv. Cryptol., pp. 129-148, 2011.
13. Shamir, "How to share a secret", Commun. ACM, vol. 22, pp. 612-613, 1979.
14. Bogdanov, S. Laur and J. Willemson, "Sharemind: A framework for fast privacy-preserving computations,"Proc. 13th Eur. Symp. Res. Comput. Security: Comput. Security, pp. 192-206, 2008.
15. R. Agrawal and R. Srikant, "Privacy-preserving data mining", ACM Sigmod Rec., vol. 29, pp. 439-450, 2000.
16. P. Zhang, Y. Tong, S. Tang and D. Yang, "Privacy preserving Naive Bayes classification", Proc. 1st Int. Conf. Adv. Data Mining Appl., pp. 744-752, 2005.
17. Evfimievski, R. Srikant, R. Agrawal and J. Gehrke, "Privacy preserving mining of association rules,"Inf. Syst., vol. 29, no. 4, pp. 343-364, 2004.